# SUPPLEMENTARY INFORMATION: Non-local means improves total-variation constrained photoacoustic image reconstruction

**Phaneendra K. Yalavarthy**[1], **Sandeep Kumar Kalva**[2], **Manojit Pramanik**[2], **Jaya Prakash**[3,*]
[1]Department of Computational and Data Sciences, Indian Institute of Science, Bangalore, 560012.
[2]School of Chemical and Biomedical Engineering, Nanyang Technological University, Singapore, 637459.
[3]Department of Instrumentation and Applied Physics, Indian Institute of Science, Bangalore, 560012.
[*]Corresponding author: jayap@iisc.ac.in

This document provides the supplementary information for the original manuscript.

## 1 Figure of merit

The Pearson correlation (PC) coefficient can be computed as,

$$PC(x_{target}, x_{recon}) = \frac{cov(x_{target}, x_{recon})}{s(x_{target})s(x_{recon})} \tag{1}$$

where, $x_{target}$ is the expected initial pressure distribution and $x_{recon}$ is the reconstructed initial pressure distribution. $cov$ denotes the covariance, and $s$ denotes the standard deviation. The CNR is defined as,

$$CNR = \frac{\mu_{roi} - \mu_{back}}{(s_{roi}^2 a_{roi} + s_{back}^2 a_{back})^{1/2}} \tag{2}$$

where, $\mu$ denotes the mean and $s$ represents the standard deviation. The subscript $roi$ and $back$ represent the region of interest and the background correspondingly in the reconstructed image. The area ratio is represented as $a_{roi} = \frac{A_{roi}}{A_{total}}$ and $a_{back} = \frac{A_{back}}{A_{total}}$. Note that in simulation studies we are aware of the exact ROI and background pixels, and the same was used to compute the CNR. The SSIM between the reconstructed image ($x_x$) and ground truth ($x_y$) is computed as,

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{3}$$

where $\mu_x$, $\mu_y$ represents the mean of $x$ and $y$ respectively, $\sigma_x$, $\sigma_y$ indicates the variance of $x$ and $y$ respectively, $\sigma_{xy}$ is the covariance of $x$ and $y$, and $c_1$ and $c_2$ are constants to stabilize the division. In case of experimental data, the signal to noise ratio (SNR) was used to evaluate the performance of different reconstruction scheme as the expected distribution is unknown, the SNR is given as,

$$SNR = 20\mathbf{log}_{10}(\frac{x_{signal}}{x_{noise}}) \tag{4}$$

where, $x_{signal}$ are the pixels corresponding to the reconstructed region of interest in the PA image and $x_{noise}$ are the pixels corresponding to the background noise in the reconstructed PAT image. In the experimental case, we are not aware of the actual image, hence we have sorted the intensity of the reconstructed image and then considered the average of first 2000 highest values as the signal and the average of the last 2000 lowest values (in absolute terms) as the background noise to compute the SNR.

**Algorithm 1** *Total Variation + Non-local Means (TV-NLM)*

AIM: Estimation of $x_{TV-NLM}$

INPUT: Obtained boundary pressure data ($b$), Model Matrix ($\mathbf{A}$) of dimension $M \times N$, Regularization Parameter ($\lambda$), Guess ($x_0$).

OUTPUT: Reconstructed Initial Pressure Rise ($x_{TV-NLM}$)

INITIALIZE: Iteration Number ($iter = 0$), $lsqr_{iter}$ = 50, Maximum Iterations ($max_{iter}$) = 10, Threshold $(th) = 0.01$, Search Window $(se_{window}) = 3$, Similarity Window $(sim_{window}) = 2$, Filtering Degree $(fil)$=2 or 3

while $iter < max_{iter}$

1. $TV = |x_{0_k} - x_{0_{k-1}}|$ where $k = 1 : N$
2. $Der_{TV}$ = $\frac{1}{max(TV,th)}$; W = diag($Der_{TV}$) $Der_{TV}$ indicates the derivative of TV
3. $x_0 = min_x\left(||Ax - b||_2^2 + \lambda|x|_{TV}\right)$; $A_{mat} = \begin{pmatrix} A \\ \lambda W \end{pmatrix}$, $b_{vec} = \begin{pmatrix} b \\ \mathbf{0} \end{pmatrix}$, $x_0$ = lsqr($A_{mat}$,$b_{vec}$,$lsqr_{iter}$). $lsqr$ indicates the least squares QR solver.
4. $x_{TV-NLM}$ = $\mathbf{NLM}(x_0, se_{window}, f = sim_{window}, fil)$
   $x_0$ = padarray($x_0$,[f f],'symmetric'); Symmetric boundary condition.
   for $i = 1 + f : \sqrt{N} - f$
   for $j = 1 + f : \sqrt{N} - f$
   1. $W_1$ = $x_0(i - f : i + f, j - f : j + f)$, i.e. identify the similarity window patch, $wmax = 0$, $s_{weight} = 0$, $avg_{weight} = 0$
   2. rmin = max($i + f - se_{window}$,$f + 1$);
   3. rmax = min($i + f + se_{window}$,$\sqrt{N} + f$);
   4. smin = max($j + f - se_{window}$,$f + 1$);
   5. smax = min($j + f + se_{window}$,$\sqrt{N} + f$);
      for $r = rmin : rmax$
      for $s = smin : smax$
      1. Divide the image into patches from the search window limits defined by $se_{window}$ i.e. $x_k \in W_2(r, s)$, where $W_2(r, s) = x_0(r - f : r + f, s - f : s + f)$
      2. For each patch compute the spatial weight and mean between $W_1$ and $W_2$, i.e. $s_{weight} = s_{weight} + val$ and $avg_{weight} = avg_{weight} + val * x_0(r, s)$ where $val = exp(\frac{-Ker \times ||W_2 - W_1||_2^2}{fil^2})$; where $Ker$ is the normalized 2-D Gaussian kernel.
      3. if $val > wmax$ wmax = val; end

      end; end
   6. $s_{weight} = s_{weight} + wmax$ and $avg_{weight} = avg_{weight} + wmax * x_0(r, s)$
   7. if $s_{weight} > 0$; $x_{TV-NLM}(i, j) = avg_{weight}/s_{weight}$ else $x_{TV-NLM}(i, j) = x_0(i, j)$

   end; end
5. $x_0$ = $x_{TV-NLM}$

end

## 2 Algorithm of proposed TV-NLM

Note that the NLM implementation has the similarity window having a size of $5 \times 5$ and the search window size was set to $7 \times 7$. The intuition for setting the search window is as follows: when a pixel lies in a smooth or homogeneous region, the size of search window should be large to estimate the original pixel value. Further if a pixel belongs to a non-smooth or heterogeneous region, the size of search window should be small to preserve edges and textures effectively. In this work the search window size was set to $7 \times 7$, which can relate to the size of structures that we are trying to resolve. Similarity window is typically set to $5 \times 5$ in the denoising literature, which is known to be robust with noise and ideal for recovering details and fine structures.[1] The NLM part of the code can be found at Ref.[2]

The NLM algorithm assigns a weighted average of all pixels ($j \in I$; with I representing the image indices) in the image to the current pixel ($x_{tv-nlm}(i)$) i.e.,

$$x_{tv-nlm}(i) = \sum_{j \in I} w(i,j) x(j) \tag{5}$$

Here the weight, i.e. $w(i,j)$, was computed as,

$$w(i,j) = \frac{e^{\frac{||W_i - W_j||^2}{fil^2}}}{Z_i} \tag{6}$$

where $W_i$ and $W_j$ indicates the currently updating patch and all other patches in the imaging region respectively, and $fil$ indicates the degree of filtering. The $Z_i$ indicates normalizing constant, given by

$$Z_i = \sum_j e^{\frac{||W_i - W_j||^2}{fil^2}} \tag{7}$$

Hence while updating the pixel, we get the terms $avg_{weight}$ and $spatial_{weight}$ corresponding to the numerator and denominator in Eq. 6.

## 3 Experimental Setup

The experimental setup used for optoacoustic tomography to scan horse-hair phantom is shown in Fig. 2 of Ref. [3] and is briefly explained here. The optoacoustic tomography setup used a Q-switched Nd:YAG laser operating at 532 nm to deliver laser pulses with pulse width of 5 ns and 10 Hz repetition rate. These nanosecond laser pulses were delivered onto the sample using Four right-angle uncoated prisms (PS911, Thorlabs) and one uncoated Plano-concave lens (LC1715, Thorlabs). A powermeter was used to measure the light fluence on the phantom, and was found to be about 9 mJ/cm$^2$ ($<$ 20 mJ/cm$^2$ : ANSI safety limit[4]). Triangular horse hair phantom (black in color) was utilized for imaging (reconstruction results shown in the next section of supplementary). The dimensions of the horse hair used was approximately 10 mm (side-length) and 0.15 mm (diameter of hair). The triangular hair phantom was glued to the pipette tips adhered on acrylic slab.[5] The generated optoacoustic signal was measured using a 2.25 MHz flat ultrasound transducer (Olympus-NDT, V306-SU) of 13 mm diameter active area and $\sim$70% nominal bandwidth was rotated over $360^\circ$ around the sample for recording the optoacoustic signals. The

acquired optoacoustic signals were amplified and then filtered using a pulse amplifier (Olympus-NDT, 5072PR), and was recorded using a data acquisition (DAQ) card (GaGe, compuscope 4227) having a sampling frequency of 25 MHz. The DAQ card was synchronized with the laser illumination using a sync signal from the laser. The reconstructed optoacoustic imaging region was having 200×200 pixels corresponding to a physical size of 40 mm×40 mm. A system matrix having a dimension of 51200×40000 (51200: 512 time samples for 100 detector positions and 40000: 200×200 reconstruction grid) was used for reconstructing the image.

*In-vivo* rat brain imaging (healthy one) was performed using a pulsed laser diode (PLD) (QD-Q1924-ILO-WATER) delivering laser pulses at 2 kHz rate with approximately 107 ns pulse duration at 816 nm wavelength. The PLD laser was controlled using a laser driving unit consisting of a water-cooling system, a 12 V power supply (Voltcraft, PPS-11810), a variable power supply (EA-PS 8160-04 T) to change the laser power, and a function generator (RIGOL DG1022) to control the pulse repetition rate (up to 2 kHz). The experimental set up for *in-vivo* imaging is shown in Fig. 1 of Ref.[6] A powermeter was used to measure the light fluence on the animal head (skin), which was measured to be 0.17 mJ/cm$^2$, and was scanned for about 12 seconds which is well within the ANSI safety limits. Healthy female rats weighing 90 $\pm$ 5 g were purchased from InVivos Pte Ltd. for performing *in-vivo* optoacoustic animal brain imaging experiment. All animal experiments were carried out according to the guidelines and regulations approved by the institutional Animal Care and Use committee of Nanyang Technological University, Singapore (Animal Protocol Number ARF-SBS/ NIE-A0263). A intraperitoneal injection with a dosage of 0.2 mL/100 g of a mixture containing ketamine (85 mg/Kg) and xylazine (15mg/Kg) was given to anesthetize the mice before performing the optoacoustic experiment. The mice was placed in the optoacoustic scanner after trimming the hair on the head and hair removal using a cream. The mice was maintained under anesthesia during the entire experiment using 0.75% isoflurane gas (Medical Plus Pte Ltd., Singapore) along with oxygen (1.2 L/min), which was delivered through a nose cone with a breathing mask covering the nose and mouth of the mice. A custom made animal holder was used to place the animal in a sitting position on its abdomen and surgical tapes were used to hold the animal in this position. Then, the animal along with the animal holder was mounted on a translational stage during the experiment. Height of the animal was adjusted using this translational stage in order to adjust the scanning plane of the rat brain for obtaining tomographic optoacoustic measurements. The photograph of the lateral plane with and without skin layer is shown in Figs 3(c) and 3(f), respectively. The generated optoacoustic signal was measured using a 5 MHz flat ultrasound transducer (Olympus-NDT, V306-SU) of 13 mm diameter active area and $\sim$70% nominal bandwidth augmented with 45 degree acoustic reflector (F102, Olympus NDT) was rotated over $360^\circ$ around the sample for recording the optoacoustic signals.

## 4 Optimal Choice of Regularization Parameter

The regularization parameter was searched from a wide spectrum of real values starting from 0.00001 to 10. The optimal regularization parameters used in this work is indicated in the table below

wherein the regularization parameter was decreased within the L1-norm implementation by a factor of 0.5 at every iteration until the tolerance condition i.e. $||Ax - b||_2^2 < 10^{-4}$ was satisified

**Table 1** Optimal regularization parameter obtained corresponding to the TV, TV+NLM and L1-norm algorithms

| Phantom/Experiment | TV | TV+NLM | L1 |
|---|---|---|---|
| Blood Vessel Phantom $SNR_d$=40dB (Fig. 1) | $\lambda$=0.00001 | $\lambda$=0.00001 | - |
| Blood Vessel Phantom $SNR_d$=30dB (Fig. 1) | $\lambda$=0.0001 | $\lambda$=0.0001 | - |
| Blood Vessel Phantom $SNR_d$=20dB (Fig. 1) | $\lambda$=0.005 | $\lambda$=0.0001 | - |
| In-vivo Rat Brain (Fig. 3) | $\lambda$=0.0001 | $\lambda$=0.0001 | $\lambda$=0.5$\times$max(abs(A'*b)) |
| Horse Hair (Supplementary) | $\lambda$=0.001 | $\lambda$=0.01 | $\lambda$=0.5$\times$max(abs(A'*b)) |

or a maximum of 50 iterations have reached. Note that the parameters like similarity window, and search window used in the NLM algorithm was fixed for all the results presented in this work.

## 5 Horse-hair Phantom results

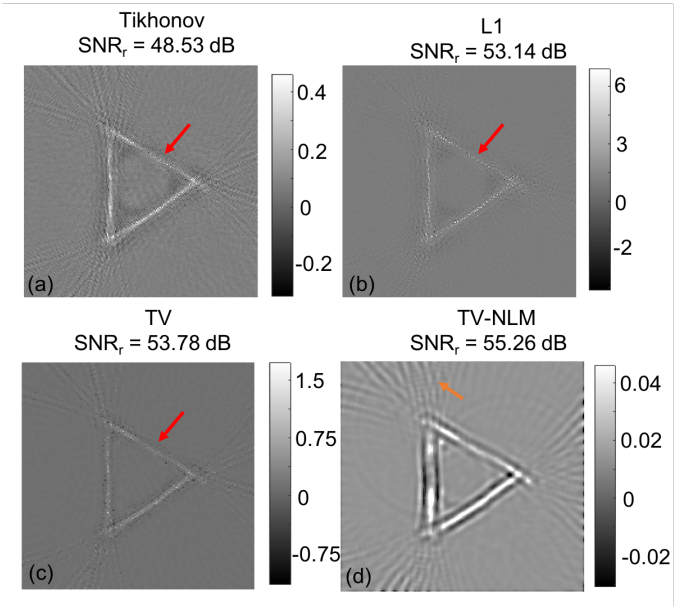


**Fig 1** Reconstructed results pertaining to horse hair phantom using (a) Tikhonov regularization, (b) L1-norm regularization (c) TV regularization, (d) proposed TV-NLM (with degree of filter=3). The SNR of these reconstructed images ($SNR_r$) was provided correspondingly at the top of each image.

The performance of standard reconstruction schemes i.e. Tikhonov, L1-norm and TV were compared against the TV-NLM method using horse-hair phantom measurements. Fig. 1(a) indicates the reconstruction results corresponding to the Tikhonov reconstruction scheme, as expected there were streaking artifacts at the edge of the reconstructed images due to limited tomographic angles. Fig. 1(b) and 1(c) shows the reconstruction results corresponding to L1-norm based regularization and TV based regularization respectively. Speckle type artifacts seems to arise with L1-norm and TV based reconstruction scheme. Moreover, all the standard reconstruction schemes seems to have discontinuity while recovering the horse hair structure which is indicated by red-arrow in Fig. 1(a)-(c). In contrary the proposed TV-NLM scheme is able to recover the horse-hair

structure more accurately (shown in Fig. 1(d)), however the streaking artifacts were still present with the TV-NLM scheme as shown by the orange arrow in Fig. 1(d). Note that introduction of NLM step with the reconstruction seems to make the image more blurrier.

*References*


1 A. Baudes, B. Coll, and J. Morel in: IEEE Conf. on Comp. Vision and Pattern Recognition, **2005**, 8624018.

2 https://in.mathworks.com/matlabcentral/fileexchange/13176-non-local-means-filter

3 N. Awasthi , S. K. Kalva , M. Pramanik, and Phaneendra K Yalavarthy, "Image Guided Filtering for Improving Photoacoustic Tomographic Image Reconstruction," *J. Biomed. Opt.*, vol. 23 no. 9, p. 091413, Sep. 2018.

4 American National Standard for Safe Use of Lasers ANSI Z136.1-2000.

5 P. K. Upputuri and M. Pramanik, "Pulsed laser diode based optoacoustic imaging of biological tissues," *Biomed. Phys. Eng. Exp.*, vol. 1 no. 4, p. 045010 Dec. 2015.

6 S. K. Kalva, P. K. Upputuri, and M. Pramanik, "High-speed, low-cost, pulsed-laser-diode-based second-generation desktop photoacoustic tomography system," Opt. Lett. vol. 44 no. 1, pp 81-84 Jan. 2019.